

Brain-inspired replay for continual learning with artificial neural networks

Year: 2020 | Citations: 540 | Authors: Gido M. van de Ven, H. Siegelmann, A. Tolias

Abstract

Artificial neural networks suffer from catastrophic forgetting. Unlike humans, when these networks are trained on something new, they rapidly forget what was learned before. In the brain, a mechanism thought to be important for protecting memories is the reactivation of neuronal activity patterns representing those memories. In artificial neural networks, such memory replay can be implemented as ‘generative replay’, which can successfully – and surprisingly efficiently – prevent catastrophic forgetting on toy examples even in a class-incremental learning scenario. However, scaling up generative replay to complicated problems with many tasks or complex inputs is challenging. We propose a new, brain-inspired variant of replay in which internal or hidden representations are replayed that are generated by the network’s own, context-modulated feedback connections. Our method achieves state-of-the-art performance on challenging continual learning benchmarks (e.g., class-incremental learning on CIFAR-100) without storing data, and it provides a novel model for replay in the brain. One challenge that faces artificial intelligence is the inability of deep neural networks to continuously learn new information without catastrophically forgetting what has been learnt before. To solve this problem, here the authors propose a replay-based algorithm for deep learning without the need to store data.